

1/8

Input file flh15625cons; Output File 15625tr
Sequence length 2286

CGTCCGTAGCTTTGAGTCCAGTGTGTTGAAGACAATCTCTGATTGTGAAGCCCTCTTTTCTCTCTCTTCTATTCTCTCT

AGAGCACTCAAGACTTTACTGACGAAAACTCAGGAAATCCTCTATCACAAAGAGGTTTGCCAATAAACTAAGACATTA

AAAGGAAAAATACCAGATGCCACTCTGCAGGCTGCAATAACTACTACTTACTGGATACATTCAAACCCCTCCAGAATCAAC

M Q A V D N L T S A P G N T 14
AGTTATCAGGTAACCAACAAGAA ATG CAA GCC GTC GAC AAC CTC ACC TCT GCG CCT GGG AAC ACC 42

S L C T R D Y K I T Q V L F P L L Y T V 34
AGT CTG TGC ACC AGA GAC TAC AAA ATC ACC CAG GTC CTC TTC CCA CTG CTC TAC ACT GTC 102

L F F V G L I T N G L A M R I F F Q I R 54
CTG TTT TTT GTT GGA CTT ATC ACA AAT GGC CTG GCG ATG AGG ATT TTC TTT CAA ATC GCG 162

S K S N F I I F L K N T V I S D L L M I 74
AGT AAA TCA AAC TTT ATT ATT TTT CTT AAG AAC ACA GTC ATT TCT GAT CTT CTC ATG ATT 222

L T F P F K I L S D A K L G T G P L R T 94
CTG ACT TTT CCA TTC AAA ATT CTT AGT GAT GCC AAA CTG GGA ACA GGA CCA CTG AGA ACT 282

F V C Q V T S V I F Y F T M Y I S I S F 114
TTT GTG TGT CAA GTT ACC TCC GTC ATA TTT TAT TTC ACA ATG TAT ATC AGT ATT TCA TTC 342

L G L I T I D R Y Q K T T R P F K T S N 134
CTG GGA CTG ATA ACT ATC GAT CGC TAC CAG AAG ACC ACC AGG CCA TTT AAA ACA TCC AAC 402

P K N L L G A K I L S V V I W A F M F L 154
CCC AAA AAT CTC TTG GGG GCT AAG ATT CTC TCT GTT GTC ATC TGG GCA TTC ATG TTC TTA 462

L S L P N M I L T N R Q P R D K N V K K 174
CTC TCT TTG CCT AAC ATG ATT CTG ACC AAC AGG CAG CCG AGA GAC AAG AAT GTG AAG AAA 522

C S F L K S E F G L V W H E I V N Y I C 194
TGC TCT TTC CTT AAA TCA GAG TTC GGT CTA GTC TGG CAT GAA ATA GTA AAT TAC ATC TGT 582

Q V I F W I N F L I V I V C Y T L I T K 214
CAA GTC ATT TTC TGG ATT AAT TTC TTA ATT GTT ATT GTA TGT TAT ACA CTC ATT ACA AAA 642

E L Y R S Y V R T R G V G K V P R K K V 234
GAA CTG TAC CGG TCA TAC GTA AGA ACG AGG GGT GTA GGT AAA GTC CCC AGG AAA AAG GTG 702

N V K V F I I I A V F F I C F V P F H F 254
AAC GTC AAA GTT TTC ATT ATC ATT GCT GTA TTC TTT ATT TGT TTT GTT CCT TTC CAT TTT 762

A R I P Y T L S Q T R D V F D C T A E N 274
GCC CGA ATT CCT TAC ACC CTG AGC CAA ACC CGG GAT GTC TTT GAC TGC ACT GCT GAA AAT 822

FIG 1A

2/8

T L F Y V K E S T L W L T S L N A C L D	294
ACT CTG TTC TAT GTG AAA GAG AGC ACT CTG TGG TTA ACT TCC TTA AAT GCA TGC CTG GAT	882
P F I Y F F L C K S F R N S L I S M L K	314
CCG TTC ATC TAT TTT TTC CTT TGC AAG TCC TTC AGA AAT TCC TTG ATA AGT ATG CTG AAG	942
C P N S A T S L S Q D N R K K E Q D G G	334
TGC CCC AAT TCT GCA ACA TCT CTG TCC CAG GAC AAT AGG AAA AAA GAA CAG GAT GGT GGT	1002
D P N E E T P M *	343
GAC CCA AAT GAA GAG ACT CCA ATG TAA	1029

ACAAATTAAC TAAGAAATATTTCAATCTCTTTGTGTTTCAGAACTCGTTAAAGCAAAGCGCTAAGTAAAAATATTAAC T
GACGAAGAAGCAACTAAGTTAATAATAATGACTCTAAAGAAACAGAAGATTACAAAAGCAATTTTCATTTACCTTTTCCA
GTATGAAAAGCTATCTTAAATATAGAAAATACTAACTGTAGCTGTATTAGCAGCAAAACAAACGACATCCAATT
GTCATGCTGCATGCAAACTACACAGAATTCATGTTTTGCAGAGTTTTGCCAAATGAGTAATCATATAATATTTACTG
TAATTTTAAAAATACATTATCGTTCACAATTTTATTTTTCATAATCAACTAAGGAAGAAGCATCAATTGGATATAATT
TCTTACCAAAAATGATAGTTAAATGTATATATATCTAGTCCCCTAACCAATCCTGACCTATTGGGATACTTATAAA
AATTTAAGTAAGTGGGATACACAAAGAATAAATACTATTAAC TTTTCATTATTAGCAAAAACCTAAGGGATTAAACTA
ATTGAACTGTATTTGATTGGACTTAATTTTTATGTTTATTTAGAAGATAAAGATTAAAGAAGACCTTTACAATAAA
GAGAAGAAATATCGAAGTCATTAAATAAGGAGACTTACTTTTATGACATTCTAATACTAAAAATATAGAAATATTTT
CTAATCTAGAGAACTAGTTTTACTAATTTTTTACAACCTCAATAATACCATCACTGACACTTACCTTTATTAATTA
GCTTCTAGAAAATAGCTGCTAATTAGGTTAATGAACATTTTACCTTAGTGAAAAAAATTAATTAAATATGATTACAAAG
TTGCACAGCATAACTACTGAGAGGAAAGTGATTGATCTGTTTGTAAATTACTTGTTTGTATTGGTGTGTATAAAATACAA
AATTTACATTAACTCTAAAAAAGGGCGG

FIG 1B

095409-10460

3/8

```
>PF00001|PF00001 7 transmembrane receptor (rhodopsin family)

Score: 184.21  Seq: 42 298  Model: 1 269
*GNI LVIWvIcRyRRMRTPMNYFIvNLAvADLLFslftMPFWMvYyvMqg
  N+L + +++++ R+ ++ + +F+ N ++DLL+ ++T+PF +++ + G
Flh15625or 42  TNGLAMRIFFQIRS-KSNFIIFLKNTVISDLLM-ILTFPFKILSDAKLG 88

      RWpFGdfmCrIWmYFDYMNMYASIFfLTcISIDRYLWAICHPMrYmRWMT
      + P+ +F+C +++ ++Y++MY SI FL +I+IDRY+ ++P++ + +
Flh15625or 89  TGPLRTFVCQVTSVIFYFTMYISISFLGLITIDRYQ-KTTRPFKTSNPKN 137

      pRHRaWvMiiIwvMSFLISMPPFLMFrWstyrDEneWNmTWCMiYDWPe
      + A++++++IW+++FL+S+P + M+ + T+R ++ N+ C++ E
Flh15625or 138 LL-GAKILSVVIWAFMFLLSLP-N-MI-L-TNRQPRDKNVKKCSF-LKSE 181

      ..wMwRwYvILmtiimgFYIPMiIMlFCYwRIYRIaRlWMRMIpswQrRR
      W +V ++ + F I ++I ++CY++I +++++ ++ +++ +
Flh15625or 182 FGLVWHEIVNYICQ-VIFWINFLIVIVCYTLITKELYRSYVRTRGVGK-- 228

      rmSmRrERRivKMliiIMvVFIIcWlPYFIvmfMDTLM.MwwFCefC.Iw
      +++++ ++II+ VF+IC+ P++ + + +TL ++ ++ +
Flh15625or 229 ----VPRKKVNVKVFIIIAVFFICFVPFHFARIPYTLSQTRDVFDCtAEN 274

      rrlWmY.IfewLaYvNCpCiNPIIY*
      +++++ ++WL ++N C++P+IY
Flh15625or 275 TLFYVKESTLWLtSLNA-CLDPFIY 298
```

FIG 2

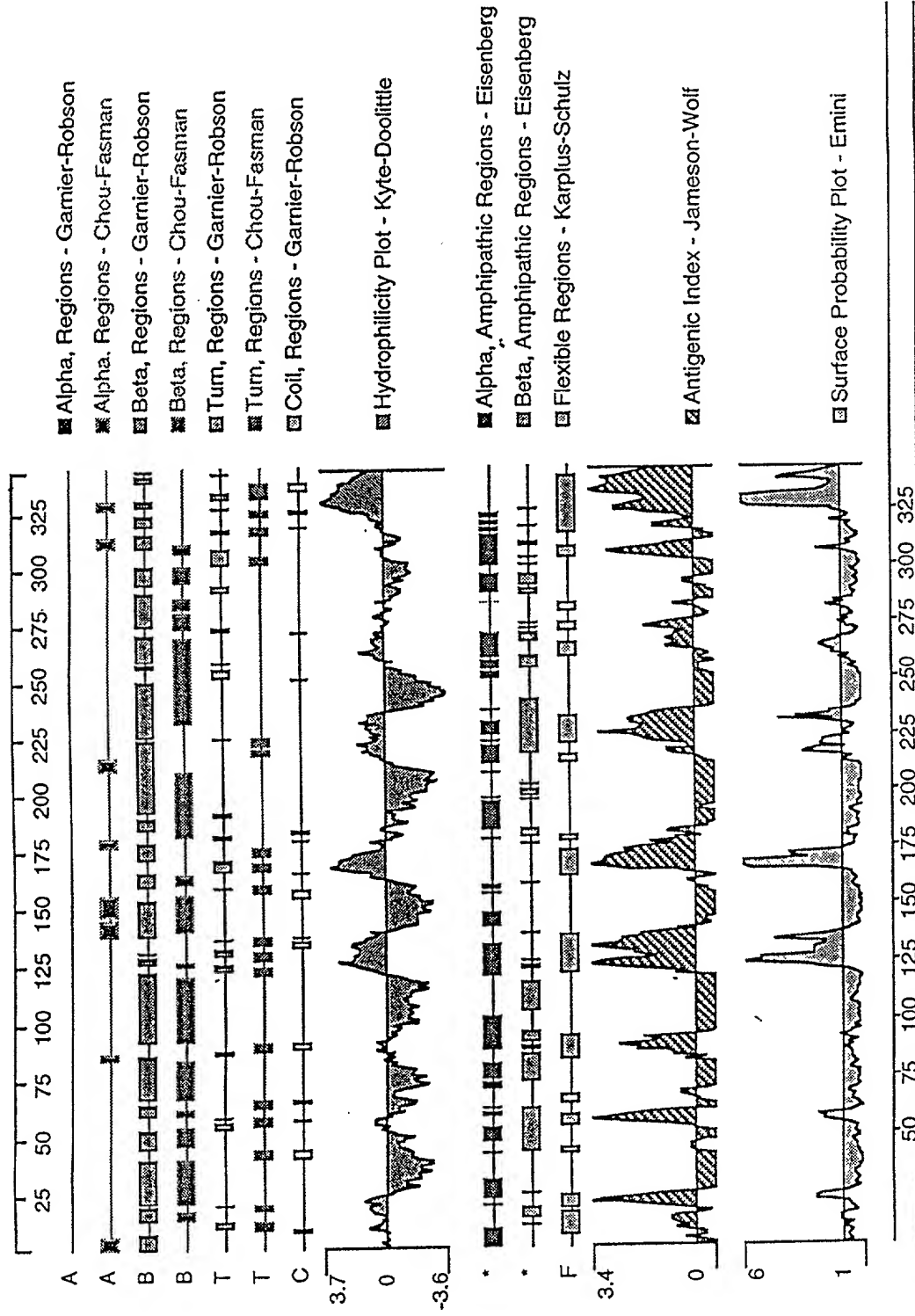


FIG 3

5/8

15625orfaa -> KD Hydrophobicity <11/1>

Protein sequence 343 a.a. MQAVDNLTSAPG ... DGGDPNEETPMZ

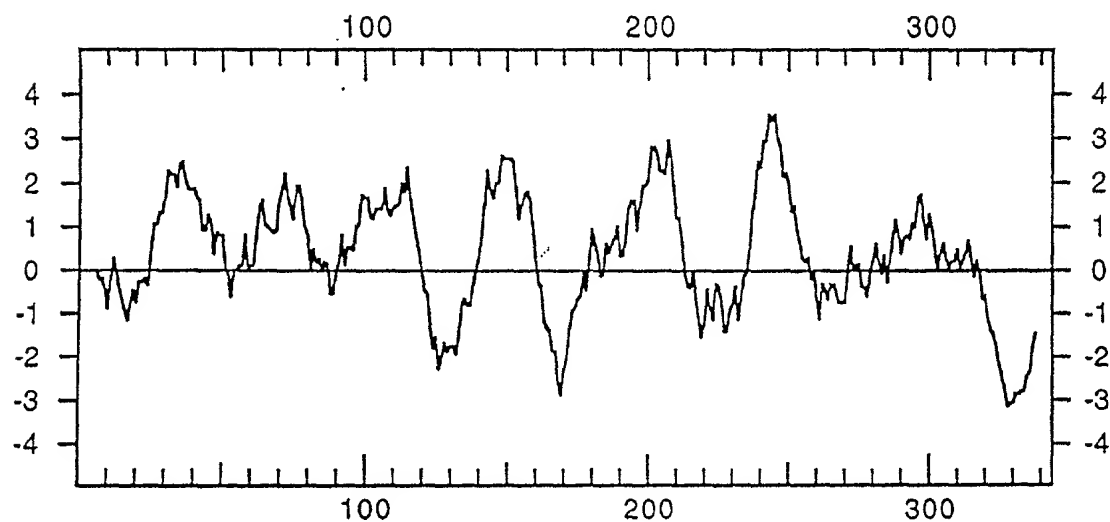


FIG 4

6/8

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.
N[^*P][ST][^*P]

Query: 6 nlts 9

Query: 13 ntal 16

>PS00004|PDOC00004|CAMP_PHOSPHO_SITE CAMP- and cGMP-dependent protein kinase
phosphorylation site.
[RK]{2}[A-Z][ST]

Query: 173 kkcs 176

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.
[ST][A-Z][RK]

Query: 126 ttr128

Query: 163 tnr165

Query: 304 sfr306

>PS00008|PDOC00008|MYRISTYL N-myristoylation site.
G[^*EDRKHPFYW][A-Z]{2}[STAGCN][^*P]

Query: 39 glitng 44

Query: 333 ggd pne 338

FIG 5

7/8

Input file 79a2cons; Output File 79a2tra
Sequence length 2272

ACGCGTCCGCAATCTCTGATTGTAAAGCCCTCTCTTCTCTCTTCTATTTCTCTATAGAACACTCAAGACTTTACTGA
TGAAAACCTCAGGAAATCTCTATCACAAAGAGGTTTGGCAACTAACTAAGACATTAAAAGGAAAATACCAGATGCCAC
TCTGCACGTTGCAATAACTACTATTTACTGGATACATTCAAATCCTCCAGAATCAACGGTTATCAGGTAACCAACAAGA

M	Q	A	I	D	N	L	T	S	A	P	G	N	T	S	L	C	T	R	19	
A	ATG	CAA	GCC	ATC	GAC	AAC	CTC	ACG	TCT	GCG	CCT	GGG	AAC	ACC	AGT	CTG	TGC	ACC	AGA	57
D	Y	K	I	T	Q	V	L	F	P	L	L	Y	T	V	L	F	F	V	G	39
GAC	TAC	AAA	ATC	ACC	CAG	GTC	CTC	TTC	CCA	CTG	CTC	TAC	ACT	GTC	CTG	TTT	TTT	GTT	GGA	117
L	I	T	N	S	L	A	M	R	I	F	F	Q	I	R	S	K	S	N	F	59
CTC	ATC	ACA	AAT	AGC	CTG	GCG	ATG	AGG	ATT	TTC	TTT	CAA	ATT	CGG	AGT	AAA	TCA	AAC	TTT	177
I	I	F	L	K	N	T	V	I	S	D	L	L	M	I	L	T	F	P	F	79
ATT	ATT	TTT	CTT	AAG	AAC	ACA	GTC	ATT	TCC	GAT	CTT	CTC	ATG	ATT	CTG	ACT	TTT	CCA	TTC	237
K	I	L	S	D	A	K	L	G	T	G	P	L	R	T	F	V	C	Q	V	99
AAA	ATT	CTT	AGT	GAT	GCC	AAA	CTG	GGA	ACA	GGA	CCA	CTG	AGA	ACT	TTT	GTG	TGT	CAA	GTT	297
T	S	V	I	F	Y	F	T	M	Y	I	S	I	S	F	L	G	L	I	T	119
ACC	TCC	GTC	ATA	TTT	TAT	TTC	ACA	ATG	TAT	ATC	AGT	ATT	TCA	TTC	CTG	GGA	CTG	ATA	ACT	357
I	D	R	Y	Q	K	T	T	R	P	F	K	T	S	N	P	K	N	L	L	139
ATC	GAT	CGC	TAC	CAG	AAG	ACC	ACC	AGG	CCA	TTT	AAA	ACA	TCC	AAC	CCC	AAA	AAT	CTC	TTG	417
G	A	K	I	L	S	V	L	I	W	A	F	M	F	L	L	S	L	P	N	159
GGG	GCT	AAG	ATT	CTC	TCT	GTT	CTC	ATC	TGG	GCA	TTC	ATG	TTC	TTA	CTC	TCT	TTG	CCT	AAC	477
M	I	L	T	N	R	R	P	R	D	K	N	V	K	K	C	S	F	L	K	179
ATG	ATT	CTG	ACT	AAC	AGG	CGG	CCA	AGA	GAC	AAG	AAT	GTG	AAG	AAA	TGC	TCT	TTC	CTT	AAA	537
S	E	F	G	L	V	W	H	E	I	V	N	Y	I	C	Q	V	I	F	W	199
TCA	GAG	TTC	GGC	CTA	GTC	TGG	CAT	GAA	ATA	GTA	AAT	TAC	ATC	TGT	CAA	GTC	ATT	TTC	TGG	597
I	N	F	L	I	V	I	V	C	Y	T	L	I	T	K	E	L	Y	R	S	219
ATT	AAT	TTC	TTA	ATT	GTC	ATT	GTA	TGT	TAC	ACA	CTC	ATT	ACA	AAA	GAA	CTG	TAC	CGG	TCA	657
Y	V	R	T	R	G	V	G	K	V	P	R	K	K	V	N	V	K	V	F	239
TAT	GTA	AGA	ACA	AGG	GGT	GTA	GGT	AAA	GTC	CCC	AGG	AAA	AAG	GTG	AAC	GTC	AAA	GTT	TTC	717
I	I	I	A	V	F	F	I	C	F	V	P	F	H	F	A	R	I	P	Y	259
ATT	ATC	ATT	GCT	GTA	TTC	TTT	ATT	TGT	TTT	GTT	CCT	TTC	CAT	TTT	GCC	CGA	ATT	CCT	TAT	777
T	L	S	Q	T	R	D	V	F	D	C	A	A	E	N	T	L	F	Y	V	279
ACC	CTG	AGC	CAA	ACC	CGG	GAT	GTC	TTT	GAC	TGC	GCC	GCT	GAA	AAT	ACT	CTG	TTC	TAT	GTG	837

FIG 6A

8/8

K	E	S	T	L	W	L	T	S	L	N	A	C	L	D	P	F	T	Y	F		299
AAA	GAG	AGT	ACT	CTG	TGG	TTA	ACT	TCC	TTA	AAT	GCA	TGC	CTG	GAT	CCG	TTC	ACC	TAT	TTT		897
F	L	C	K	S	F	R	N	S	L	I	S	M	L	K	C	P	N	S	A		319
TTC	CTT	TGC	AAG	TCC	TTC	AGA	AAT	TCC	TTG	ATA	AGT	ATG	CTG	AAG	TGC	CCC	AAT	TCT	GCA		957
T	S	Q	S	Q	D	N	R	K	K	E	Q	D	G	G	D	P	N	E	E		339
ACA	TCT	CAG	TCC	CAG	GAC	AAT	AGG	AAA	AAA	GAA	CAG	GAT	GGT	GGT	GAC	CCA	AAT	GAA	GAG		1017
T	P	M	*																		343
ACT	CCA	ATG	TAA																		1029

ACATATTAACTGAGGAAATATGTCAATCTCTTTGCGTTCAGAACTCATTAAAGCAAAGCGCTACGTAAAAATATTA
GACGAAGAAGCAACTGAGTTAATAACAATGACTCTTAAACATGTAATAGAAGATTACAAAAGCAATTTTCATTACC
TTTCCAGTATGAAAAGCTATGTTAAATATAGAAAATACTAATCTAACCTGTAGCTGTATAGTATCAAAACAAATGACATC
CAATTGGCATGCTGTCATGCAAACTACACAGAATTCACGTTTTGCAGAGTTTTGCCAAATGAGTAATCATATAATATC
TACCGTAATGTTTAAATACATTATTGCTCACGATTTTATTCTTCATAATCAACTAAGGAAGAATTATCAATTGGATA
CAATCTTCTTACAAAAATGACACTTAAATGTATATATATCTAGCCCCTAACCAATCCTGACCTATTGGGATACTT
ATAAAATTTGAGTAAGTGGGATACACAAAGAATAAATACTATTAACCTTTTAATTATGAGCAAAACCTAAGGGTTAA
TTTAACTAATTGAACTGTATTTGATTGGACTTAATTTTTTGTATTATTAAGAAGACACTGAAGAAGACCTTTACAA
TAAAGAGAAGAAATATCAAAGTCATTAAATAAGGAGAGTTACTTTTATGATATTCTAACACTAAACAATATAGAAATA
TTTCCTTAATATTAGTTTCTAGAGAACTAGTTTTACTAATTTTTTACAACCTCAATAATACCATCATTGACACTTACC
TTTATTAAC TAGCTTCTAGAAAATACCTGCTAATTAGGTTAATGAACATTTTATGTTAGTGAAAAAATTAATTAAATA
TGATTACAAAGTTGCACAGCATAACTACTGAAAGTGATTGATCCATTTGTAATTATTTGTTTGTACTGGTGTGTATAAA
ATACAAATTTACATTAACTCTAATCACCAAAAAAAAAAAAAAAAAAAGGGCGG

FIG 6B

0954093 0954094